DOCUMENT RESUME

ED 127 532 CG 010 760

AUTHOR Sedlacek, William E.
TITLE Recent Developments in Test Bias Research. University of Maryland Cultural Study Center Research Report No. 2-76.
INSTITUTION Maryland Univ., College Park. Cultural Study Center.
PUB DATE 76
NOTE 12p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Attitudes; *Bias; *Culture Free Tests; Decision Making; Models; *Racism; *Sex Discrimination; *Social Change; Social Discrimination; State of the Art Reviews; *Test Bias

ABSTRACT

Three types of test bias are discussed: content bias, atmosphere bias, and use bias. While much concern has been expressed about the content of tests or the atmosphere in which they are given, it is more important to focus on how tests are used in making decisions about people. Four models of test use are defined: regression model, constant ratio model, conditional probability model, and the equal probability model. The writer feels that, rather than asking if tests are biased, one should ask if the society is biased, since tests are always employed in a cultural context. The answer to eliminating test bias is to eliminate racism and sexism in society. A six-stage model to eliminate racism and sexism is presented: (1) Cultural, Racial and Sexual Differences; (2) How Racism Operates; (3) Examining Racial and Sexual Attitudes; (4) Sources of Racial Attitudes; (5) Changing Behavior: What Can Be Done?; and (6) Changing Behavior: How It Can Be Done. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

CULTURAL STUDY CENTER

DATA CENTER

UNIVERSITY OF MARYLAND

College Park, Maryland

ED127532

RECENT DEVELOPMENTS IN TEST BIAS RESEARCH

William L. Sedlacek

Research Report # 8-76

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

CULTURAL STUDY CENTER
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND

RECENT DEVELOPMENTS IN TEST BIAS RESEARCH

William E. Sedlacek

Research Report # 2-76

Summary

Three types of test bias are discussed: content bias, atmosphere bias, and use bias. While much concern has been expressed about the content of tests or the atmosphere in which they are given, it is more important to focus on how tests are used in making decisions about people. Four models of test use are defined: regression model, constant ratio model, conditional probability model, and the equal probability model. The writer feels that, rather than asking if tests are biased, one should ask if the society is biased, since tests are always employed in a cultural context. The answer to eliminating test bias is to eliminate racism and sexism in society. A six-stage model to eliminate racism and sexism is presented: I Cultural, Racial and Sexual Differences, II How Racism Operates, III Examining Racial and Sexual Attitudes, IV Sources of Racial Attitudes, V Changing Behavior: What Can Be Done?, VI Changing Behavior: How It Can Be Done.

The word "test bias" has been with us for some time. However, only recently has it been clear that the concept is complex and cannot be glibly passed over without some analysis and specification. Flaugher (1970, 1974) discussed three general approaches to test bias or unfairness: content bias, atmosphere bias, and use bias. The purpose of this article will be to discuss some specific points under these three general categories, and ultimately to offer a solution to test bias.

Content Bias

Much of the initial concern about test bias focused on the content of the items on the tests, or even the instructions. The language on most tests is oriented toward white, middle class males. Cries of unfairness sprang from many quarters in the 1960's and 1970's . This has led to the development of tests geared to other cultural, racial, and sexual groups. For instance, the Dove Counterbalance Intelligence Test, or so-called "Chitlin'" test was developed in the mid 1960's to assess the intellectual capacities of street-wise, west coast blacks (see Sedlacek and Brooks, 1976). Additionally, Williams (1975) has developed the BITCH (Black Intelligence Test of Cultural Homogeneity) Test as a culture specific test tapping black experience.

Many instruments have been revised in recent years to eliminate sexist wording and content. The most well known and widely used of these is likely the Strong-Campbell Interest Inventory (Campbell, 1974).

The biggest problem with focusing on test content in understanding test bias is that we do not get any information on whether the content makes any difference. It is not clear whether it results in different or incorrect scores for cultural-racial minorities and women. Additionally, and more importantly, it is not clear whether the content results in different decisions being made, or outcomes to those taking the tests.

Boyd (1975) provides us with a good example of the limitations of focusing on content in attempting to understand test bias. She revised Holland's Self-Directed Search to eliminate all sexist referents such as fireman, salesman, etc. She gave the nonsexist form to one group of females and the unaltered form to another group and found no differences in interest profiles or the occupational choices made. Thus the altering of content made no difference in outcome.

Atmosphere Bias

Several studies have shown that the characteristics of test administration can adversely affect the test performance and motivation of cultural-racial minorities. For instance, Katz (1970) and Epps (1974) found that the race of the examiner and the examinee's perceptions of the use of the test can affect minority student performance . As might be expected, white examiners and the potential for negative consequences resulting from the test tended to lower minority student scores. Breland et al. (1974) also found that passages in a test by and about black people made for slight increases in black examinee scores. However, the Breland study may be more related to content than atmosphere.

There is no current evidence that atmosphere bias operates with female examinees.

Use Bias

The unfair way tests have been used has received increased attention in recent years as perhaps the best way to approach the subject of test bias. Here the emphasis is on outcomes; what happens to cultural or racial minorities or women as a result of taking tests. How are counselors, admissions officers and educators using the information as they make decisions?

A full treatment of this question would involve many complex statistical and measurement arguments beyond the scope of this article. However, several test bias

models will be briefly noted.

Regression Model: If we consistently make predictions for a given group, using a regression equation, which are too low or too high for a particular group, we have biased predictors or tests (Cleary, 1968).

Constant Ratio Model: Thorndike (1971) proposes that a test is fair if applicants are selected so that the ratio of the proportion selected to the proportion successful is the same in all subgroups.

Conditional Probability Model: Cole (1973) argues that all applicants who, if selected, are capable of being successful should be guaranteed an equal opportunity to be selected, regardless of their group membership.

Equal Probability Model: All applicants who are selected should be given an equal chance of being successful, regardless of group membership (Linn, 1973).

Petersen and Novick (1976) also present the converse of the last three models. That is, they forcus on those rejected rather than those selected. The models presented in this section of the article apply most readily to situations in education or employment involving selection but could be applied to any circumstance where more desirable and less desirable outcomes can be specified. For further discussion of these and related issues in "Use" models of test bias, see Hanson et al. (1973), Linn (1973) and Petersen and Novick (1976).

The Elimination of Test Bias

The definition of test bias preferred by the writer is a general one which says that if the results of a given test result in negative outcomes for a given person because that person is identified with a certain group, we have test bias. The group may be cultural, racial, sexual, ideological, etc. This definition suggests that we are asking the wrong question if we say "Is this test biased?" A test is designed to reflect a certain ability, trait, etc. in a given cultural context.

The test items will reflect certain aspects of our society and the test will be interpreted and acted upon by counselors, researchers, parents, students, etc. in ways that reflect a culture. So a test does not exist as an independent entity.

So the appropriate question to ask is:"Is the society biased?" The answer is yes. If our society is biased (racist, sexist, etc.) then our tests will be also. Thus it is inappropriate to focus so much of our attention on tests. We cannot make tests unbiased unless we eliminate bias in society. If we were to totally abandon the use of tests today, would we eliminate any serious problems? I think not. Counselors, teachers and administrators would still act in the same way and make the same decisions with other, perhaps more subjective data. At least with tests we have something we can study and analyze to perhaps better understand our biases.

It is suggested that we focus on the general term racism and how to eliminate it as the real solution to test bias. Thus to further develop our test bias definition above, we can define individual and institutional racism (Sedlacek, 1974; Sedlacek and Brooks, 1976). Individual racism is an action taken by one individual toward another which results in negative outcomes, because the other person is identified with a certain group. The group may be racial, cultural, sexual, ideological, etc. Institutional racism is action taken by a social system or institution which results in negative outcomes for members of a certain group or groups. Research indicates that we have developed a system of education, which involves the use of tests, which tends to work against cultural minorities (blacks, Chicanos, Native Americans, etc.) and women (Sedlacek, 1974; Sedlacek and Brooks, 1976). Decisions are constantly made that tend to work for middle class, suburban, white males and against other groups.

What can be done about racism? Sedlacek (1974) and Sedlacek and Brooks (1976) have developed a practical six-stage model or approach to the elimination of racism.

The stages are progressive in that it is important to work through them in order. The model has been employed in a variety of educational and noneducation settings and is typically done in a workshop context. A brief outline of the six stages and some of the key points covered in each follows:

Stage I - Cultural, Racial and Sexual Differences

(1) Cultural, racial and sexual differences exist and they should be openly discussed and understood by all. (2) Differences should be approached and presented positively, in and out of the classroom. (3) Expressions of racial, cultural and sexual identity are necessary and healthy for cultural and racial minorities, women, and also for the rest of society.

Stage II - How Racism Operates

Aside from the definition of racism presented above, a few other points covered in this stage include:

(1) The definitions of racism are <u>behavioral</u> in that <u>results, not intentions</u>, are important. Most racism is unknowing or unintentional.

(2) Those who have the power to influence others, whether or not they are in the numerical majority, are the primary perpetuators of racism.

Of the many possible examples of racism which could be presented, a few are:

(1) Teachers have low expectations of minority student performance.

(2) Counselors tend to direct women toward social science and education, and away from the physical sciences.

(3) Relatively few minorities or women are in supervisory or central staff positions in education.

Stage III - Examining Racial and Sexual Attitudes

(1) Most people have stereotyped or negative attitudes toward other racial,

cultural or sexual groups.

(2) Racial and sexual attitudes may directly influence behavior.

(3) Racial and sexual attitudes can be measured and analyzed. The Situational Attitude Scale (SAS) (Sedlacek and Brooks, 1970, 1972) and the Situational Attitude Scale- Women (SASW) (Herman and Sedlacek, 1973) were designed for this purpose.

Stage IV - Sources of Racial Attitudes

(1) Textbooks and tests help perpetuate racial and sexual stereotypes.

(2) The nature of prejudice and racism should be taught at all educational levels. This is particularly crucial for young whites because they are not likely to avoid becoming racists without help.

(3) The study and consideration of prejudice and racism should be an integral part of the regular academic and nonacademic programs and not isolated as a one-occasion experience.

Stage V - Changing Behavior: What Can Be Done?

(1) Goals must be stated so as to provide direction for change.

(2) Goals should be as specific and operational as possible.

(3) Strategies are separate from goals in that they are ways of accomplishing goals.

(4) All goals must be evaluated as to their extent of accomplishment.

Stage VI - Changing Behavior: How it Can be Done.

(1) The only test of a strategy is whether it works.

(2) Most limitations to effective change are self imposed.

(3) More people are capable of affecting institutions than ever give it a try.

(4) The effective change agent is irreverent toward most social institutions.

A successful implementation of the model requires much more detailed information on content, techniques to employ at various stages, and the judicious employment of racial and sexual roles to help make points,which are detailed in other sources (Sedlacek, 1974; Sedlacek and Brooks, 1976).

The final thoughts I would like to leave with you are the necessity of taking some action against racism and the difficulty of change. It is always easier and more comfortable to go about our business in the usual way, to read a few articles and blame something like tests for our problems. I regularly think of a quote from the *Pogo* comic strip to boost my motivation: "We have met the enemy, and they is us!"

References

Boyd, V.S. The linguistic structure of the Self-Directed Search: A study in sex-role stereotyping. Unpublished doctoral dissertation, University of Maryland College Park, 1975.

Breland, H.M., Stocking, M., Pinchak, B.M., & Abrams, N. *The cross-cultural stability of mental test items.* Project Report 74-2. Princeton, New Jersey: Educational Testing Service, 1974.

Campbell, D.P. *Manual for the Strong-Campbell Interest Blank.* Stanford, California: Stanford University Press, 1974.

Cleary, T.A. Test bias: Prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement,* 1968, 5, 115-124.

Cole, N.S. Bias in selection. *Journal of Educational Measurement,* 1973, 10, 237-255.

Epps, E.G. Situational effects in testing. In Miller, L.P.(ed.), *The testing of black students: A symposium.* Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1974.

Flaugher, R.L. *Testing practices, minority groups, and higher education: A review and discussion of the research.* Research Bulletin 70-41. Princeton, New Jersey: Educational Testing Service, 1970.

Flaugher, R.L. Some points of confusion in discussing the testing of black students. In Miller, L.P. (ed.), *The testing of black students: A symposium.* Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1974.

Hanson, G.R., Belcher, L., Sedlacek, W.E., & Thrush, R. Alternatives to a moratorium on testing: An ACPA Commission IX position paper on test bias in the college setting. Washington, D.C.: American College Personnel Association, 1973.

Herman, M.H. & Sedlacek, W.E. Sexist attitudes among male university students. *Journal of College Student Personnel,* 1973, 14, # 6, 544-548.

Katz, I. Experimental studies of Negro-white relationships. In Berkowitz, L. (ed.). *Advances in experimental social psychology.* Vol. 5. New York: Academic Press, 1970.

Linn , R.L. Fair test use in selection. *Review of Educational Research,* 1973, 43, 139-161.

Petersen, N.S. & Novick, M.R. An evaluation of some models for culture-fair selection. *Journal of Educational Measurement,* 1976. 13. 3-29.

Sedlacek, W.E. Racism in society: A behavioral model for change. Behavioral Sciences Tape Library #82220. Leonia, New Jersey: Sigma Information, Inc., 1974.

Sedlacek, W.E. & Brooks, G.C., Jr. Measuring racial attitudes in a situational context. *Psychological Reports,* 1970, 27, 971-980.

References (continued)

Sedlacek, W.E. & Brooks, G.C., Jr. *Situational Attitude Scale (SAS) Manual.* Chicago: Natresources, Inc., 1972.

Sedlacek, W.E. & Brooks, G.C., Jr. *Racism in American education: A model for change.* Chicago: Nelson Hall, 1976.

Thorndike, R.L. Concepts of culture-fairness. *Journal of Educational Measurement,* 1971, 8, 63-70.

Williams, R.L. The BITCH-100: A culture specific test. *Journal of Afro-American Issues,* 1975, 3, 103-116.